INFORMATION STORAGE AND RETRIEVAL
IN
ADMISSIONS AND REGISTRARS OFFICES

By Artel Ricks

Director, Paperwork Standards and Automation Division
National Archives and Records Service, GSA

CONTENTS

AMERICAN ASSOCIATION OF COLLEGIATE REGISTRARS AND ADMISSIONS OFFICERS

Prepared for the 53rd AACRAO Annual Meeting,
April 1967, Denver, Colorado, under the auspices of the
Committee on New Developments and Techniques

Committee Members

Chairman, Verner Ekstrom, Director of Admissions and Records, Colorado State University
John F. Chaney, Director of Statistical Service Unit, Office of the Provost, University of Illinois
Warren Haffner, Associate Registrar, Pennsylvania State University
M. D. Scherer, Registrar, Indiana University
William E. Slaby, Registrar, Wayne State University
Orrin T. Wheeler, Assistant Registrar, Xavier University

INFORMATION STORAGE AND RETRIEVAL

IN

ADMISSIONS AND REGISTRARS OFFICES

By Artel Ricks

You walk to the file and search for a student's application for admission and you have performed an exercise in information retrieval. You turn a little knob and a roll of microfilm speeds past you at blizzard speed and comes to a halt at the desired spot where you find the grade given some complaining student in a mathematics class completed last semester. You request a list of seniors who can speak Spanish and the computer prints it out. You seat yourself at a console and in seconds you see on a TV screen before you a student's permanent record which had been updated electronically on videotape only moments before. With the touch of a button you have a transcript on paper.

These examples may seem poles apart but they are all within the spectrum of "information storage and retrieval" when the term is loosely applied.

It is within this broad concept of information retrieval that your committee on New Developments and Techniques has asked me to address my remarks. So let's begin with the storage and retrieval of hard copy--"paper" as we used to call it.

HARD COPY STORAGE AND RETRIEVAL

The data assembled and filed by admissions and registrars offices might be classified as entrance data, student description data, historical academic data, and statistical data. Every office will have its own groupings of files and arrangements under which to store this data, but this paper assumes that the following five files are typical:

1. Admissions Files. Include applications for admissions and all supporting papers and evaluation work sheets. One file folder for each applicant. May be consolidated with registration papers upon registration of the applicants.

2. Admissions Reference Card Files. Contain limited entrance and student identification information.

3. Registration Case Files. Contain most or all of the admission papers, supplemented by registration and other information about the student since his enrollment, other than his permanent record. One file folder for each applicant.

4. Permanent Record Files. Historical record of courses taken and grades received. One or several letter size sheets per student serving as the source of transcript information.

5. Statistical Files. Computer print-outs or other statistical tabulations and reports, filed by subject.

Of course there will be other files such as registration card files, general administrative subject files, class rosters, final class lists, and others. But the five are presumed to be the most typical and basic of the records maintained, and it is in this context that the comments in this paper are made.

Misfiles

Audits have shown that the typical office struggles along with nearly 3 percent of its papers or folders physically misfiled. This is too high. How much lower the figure should be depends on the kind of file involved and the impact of "can't finds" and delayed findings. For the files in question, there is little excuse for misfiles of more than 1/2 of 1 percent. If it is higher, check first to see if you have fixed responsibility on someone for the files and barred all others from pulling or replacing folders or papers in the file. Too many cooks spoil the broth in filing as well as in cooking. Misfiles can also be substantially reduced and retrieval made more certain by a variety of techniques we will discuss.

Time Required

In most subject files, it should normally take no more than 2 to 3 minutes to locate any item requested from the file. In case or card files arranged in name or number sequence, the figure should be scaled down to less than a minute depending on the proximity and size of the file and the finding aids employed. Parenthetically, we might note here that only the most expensive automated systems are faster than manual retrieval systems when the object is to find a folder by name or number, if the name or number is known at the outset.

Filing Arrangements

The most common filing arrangements are by name, number, or subject. The rule for selection is simple--file by the feature most often known when the records are needed. If the users are more likely to know the student's name than his identification number, then file by surname. To take care of those instances, if numerous, in which the number is known but not the exact name, you may need an "index" in number sequence.

Our rule, to file by the feature most often known, is simple and logical, but it is surprising how many people have been persuaded to file by number because they have heard that "it is faster and more

accurate" even though this might force them to find their way to the desired record by way of name indexes. I know of one instance in which it was necessary to work one's way progressively through four indexes, one leading to the other, to find the number under which the records were filed for an individual. Yet, 99 percent of the time they started out with the name. Nor were there any offsetting advantages.

Alphabetical filing and retrieval, performed manually, is nearly as fast as numerical for the kinds of records we are talking about if the name is known at the outset; it is much faster than numerical filing and retrieval if we must refer to an index. Alphabetical filing and retrieval, performed manually, is also nearly as accurate as numerical filing; it is usually more accurate than a numerical arrangement that is dependent upon the accuracy of the index file as well as the accuracy of the numerical file.

The fact that computers can sort numerical information more efficiently than alphabetical does not change our rule insofar as the basic hard copy records of the office are concerned. There are exceptions to be sure, but the considerations should be carefully weighed if records are to be filed by any feature other than that most often known when the records are needed.

Name Filing

Since names are the feature most often known for most of the files in question, we will give particular attention to the special problems that name-filing arrangements entail.

How would you alphabetize such names as these?

d'Alembert	MacDonald
DiStefano	McAllister
El-Abd	O'Donnel
El Al	VonRath

The problem is to insure consistency. Only if standard rules have been followed can we be assured of retrieving records in an alphabetic name file, particularly if it contains over a thousand names. We need rules for prefixes and hyphenated surnames. We need rules for filing and finding married women's names, abbreviated first names, nicknames, titles, and abbreviated designations such as Sr., Jr., and others. One excellent source for a practical list of rules is the publication, "Rules for Alphabetical Filing," published by the American Records Management Association. (738 Builders Exchange, Minneapolis, Minnesota; price $4.00)

Guides for Name Files

"Guide cards" with labels that protrude above the file folders are essential for speed and accuracy in both filing and retrieving

records. Yet they are usually either absent altogether or poorly distributed throughout the file. One guide card for every 10 to 15 folders is considered a good standard.

Name files present a seemingly difficult problem in achieving an equal spacing of guide cards throughout the file because surnames are simply not evenly distributed throughout the alphabet. In fact, six letters of the alphabet (B, C, H, M, S, and W) contain over half the names of the people that live in the United States. If your name file is not large, the easiest way to meet the standard and insure an even distribution is to buy guide cards that permit replacement of the labels. Then, go through your file and develop your own labels for every 10 or so folders. If your file is very large, say over 10,000 names, it might be preferable to turn to a file equipment company for professionally developed alphabetic guide labels which come in sets tailored to different volumes of records. Let me caution you, however, that these standard sets must be modified if you have a heavy population of foreign students.

Guide cards, because they receive heavy use and are the main guideposts to the folders, should be visible, rugged, and reusable. Angular tabs are easiest to read. Heavy metal tabs on pressboard assure a long life. Replaceable labels will permit reuse of the guide cards for many years. The cost figures resulting from our experience are those charged to Government, and you will likely pay at least 50 percent more.

Phonetic Filing

We are frequently asked about "soundex" for name filing as a means of assuring more certain retrieval. Under this system, similar sounding names such as Johnsen, Johnston, Johnstone, and Jonston, would all be filed together under an alpha-numeric code. It is rapid, accurate, and easy to learn. It has been used successfully for years in the 200-million name file of the Social Security Administration in Baltimore. It is of particular worth in a situation in which the exact spelling of surnames is frequently unknown and difficult to verify. If that is a significant problem in your office, I suggest you contact your local office of the Remington-Rand company for further information. Records R

Folders

We have found very little need for special file folders as a device for speeding filing and finding despite all salesmanship to the contrary. We prefer a folder that is straight-cut but reinforced across the top for the following reasons:

1. Those with narrow protruding tabs tend to become dog-eared and difficult to read.

2. The eye can locate desired folders much faster if the labels are in one straight row rather than zigzagged across the drawer.

3. Guide cards, since they extend above the tops of the folders, are much better guideposts for subdividing the file than staggered folder tabs.

Our recommended all-purpose folder is kraft, 11 points in thickness. For records that will be used only a year or two, or only occasionally, the manila 9.5 point folder is satisfactory and cheaper. The heavy duty folders are for unusually long and rugged usage beyond that typically found in admissions and registrars offices. Folder labels that attract the eye and contain clear type or print are essential for accurate and speedy filing and retrieval. We now have available a label that is easy to prepare on the typewriter and can be applied quickly by peeling off the paper protecting the pressure-sensitive adhesive on the back. No moistening is necessary.

Filing Cabinets

American ingenuity has devised almost every kind of storage device imaginable for every size paper and for every unusual filing problem. We will consider here only the ones most likely to be pertinent to the typical records in your offices.

To begin with, the workhorse of the filing cabinet field is the 5-drawer cabinet. A drawer usually holds about 4,000 sheets, 150 folders, and 15 to 20 guides.

Fire-insulated files are very expensive but might be justified for such papers as the Permanent Record if other provision has not been made for the reconstruction of the records if destroyed. Use of insulated files should be confined to non-fire resistant buildings or areas.

It is usually wasteful to purchase theft-resistant files with combination or other special locks. For most needs, restricted access during the day and a locked door at night will serve.

Shelf Files

Shelf filing is becoming increasingly popular. If we use shelving that is seven shelves high, we can store the same linear footage of records in about half the space required by 4-drawer filing cabinets. One reason is that shelving requires only a 30-inch aisle space as against 36 inches for filing cabinets.

We prefer the less expensive shelf files which cost about one-half as much as file cabinets per linear foot of storage capacity. However, the special guide cards and folders that should be used cost about twice those used in cabinets, and this can offset the purchase advantage of shelving within a few years.

The advantages of speed and accuracy in filing and retrieving records using shelf filing is a subject being continuously debated.

The detractors of shelf filing point out that it is harder to reach records on the top and bottom shelves and that side labels for guides and folders are harder to prepare and read. Those who favor shelf filing claim that studies show that inserting and retrieving the records is about 25 percent faster with shelf filing and that this more than offsets the time required to prepare the folders and guides at the outset if the records are quite active. We favor the use of shelf files.

In purchasing shelf filing, be wary of the kind with doors on the shelves. If the doors must be opened every time the files are used, we lose some advantage of speed. No doubt this accounts for the fact that where shelf files with doors have been purchased, you are likely to find the doors left constantly open and the extra money they cost wasted.

It seldom pays to convert an existing collection of records to shelf filing because of the costs entailed. It is most likely to pay if applied to new files such as you set up for new admissions applications and registrations every year.

Mechanized Files

Mechanized files are used extensively in our offices today. This type of equipment contains shelves which are electrically moved to the desired working position by means of push buttons. The horizontal type is generally used for card reference files while the upright type is more typically for folder files.

They look modern and efficient and are a product of prosperous times. But they have proven an albatross around the neck of the office manager who purchased them because they looked efficient rather than because they filled a proven need. They are relatively immobile, particularly the upright variety. Usually only one person can work at the file at a time. Thus it does little good to have large quantities of reference cards stored in a mechanized rotary file if the number of requests for information during rush periods exceeds the ability of one clerk.

While mechanized files can reduce physical strain because they bring the records within easy reach of the user, they are not necessarily faster than well arranged tub or shelf files for the same volume of records.

Mechanized files generally cost about 10 times that of manual files. Hence, to justify their purchase, the records must be very active but not more so than one clerk per file unit can handle. Some Government agencies have adopted the standard that such equipment should be purchased only if the cost can be amortized by the benefits within 18 months.

Preservation and Disposal

Any discussion of hard copy storage and retrieval would be incomplete without considering what to do with noncurrent records. I understand that AACRAO has a booklet on the subject called "Retention of Records."

Avoid screening if possible! By screening we mean going through the file to separate those papers to be retained from those eligible for disposal. Almost inevitably the screening process is extremely costly, seldom less than $5 to $6 per filing drawer. At best, screening costs so much that it is worth doing only if (1) over half of the papers to be screened can thereby be destroyed, and if (2) failure to screen means permanent retention of a high percentage of valueless material.

The proper time to avoid screening is when the files are created. If records of temporary value are filed separately from those of more permanent value, screening will not be necessary. Sometimes this is accomplished by the simple method of putting the records that can be destroyed after a relatively short time on the left side of a folder and on the right side attaching papers that must be kept for years. When the temporary papers have served their purpose, they can be removed quickly from the folders.

As I view your records, most of them do not have a long, useful life. I say this despite the response to a question in your summer meeting of 1963--"How long are the application, correspondence, advanced standing, etc. retained?" Of the 56 institutions represented, 3 said for 1 year; 3 indicated destruction at graduation; 18 destroyed after 5 years; 1 after 10 years; and 31 reported that they keep all information permanently. If I am correct that these and most other admissions and registrars office records should be short-lived, then screening should not be necessary. The entire contents of the folders should be kept intact and destroyed "after 5 years" or after some other specified period.

Records that must be kept in the office for a school year and thereafter will be used only occasionally should ordinarily be moved out of the office at the end of the year. Why not put them in boxes, give the boxes consecutive numbers, prepare a list of the contents of each box (e.g., names on the first and last folders in each box), and store them on shelves in a records center located in less expensive space? When a record is needed, a clerk may determine the appropriate box number by referring to the list prepared for this purpose.

We do this in the Government. Records referred to less than once a month per linear foot of records are sent to Federal Records Centers, sometimes hundreds of miles away. An agency can get back any record within 24 hours or, if sufficiently urgent, the information desired can be furnished by telephone. We are able to hold 1 cubic foot of records in such centers for 1/12 of what it would cost in offices.

Before leaving this subject, I think it should be clear that I am not suggesting the use of the kind of "records center" that is

located in the corner of the administration building basement, where inactive records have been dumped in one horrible heap, and to which no one will go if there is any possible alternative. To the contrary, I mean a clean and well lighted room with adequate shelving and with the boxes of records arranged in an orderly and easily accessible way. A good records center involves much more. You may find the publication, Federal Records Centers, helpful. (Available from the U. S. Government Printing Office, Washington, D. C., for 20 cents.)

MICROFILM STORAGE AND RETRIEVAL

Five major uses are being made of microfilm today. To:

1. Save space
2. Guard against loss of vital information
3. Preserve valuable, irreplaceable records
4. Reproduce and distribute documents
5. Improve retrieval of information

The first use of microfilming, as a space saving technique, was greatly overdone in the past because in many instances it would have been cheaper to store the records in records-center space in a separate area. This is particularly true of case folders containing papers of various sizes, thicknesses, and inscriptions, some of which are stapled together or fastened down. Studies have shown that such records can be stored for an average of about 40 years in a center before the cost would equal that required to microfilm them.

The second use of microfilming, guarding against loss of vital information such as that essential to resumption of operations following a disaster, is usually a valid justification for a limited group of records. Of course, a copy of this microfilm will be stored at a separate location outside of any likely disaster area. From the minutes of a previous meeting of your Association, I note that security is apparently considered as the major reason you have microfilmed the permanent records.

The third use, microfilming to preserve valuable irreplaceable records, pertains to situations where there is danger of their becoming lost or damaged through constant use. Thus far we have not been able to assign a life expectancy figure to microfilm. We believe that it will be usable after perhaps 100 years, given proper care, but at the same time we are not destroying records that have been microfilmed which have permanent reference value.

The fourth use, microfilming for reproduction and distribution of documents, is becoming increasingly important. It can result in

substantial savings in funds and elapsed time in situations where the number of copies needed is not sufficient to justify use of conventional printing and distribution methods, or where there is a long-term unpredictable need for an occasional copy of individual documents.

This brings us to the fifth and, perhaps, most important use of microfilm--for improving retrieval of information. In the past, microfilming was usually justified on the basis of one or more of the preceding four uses, but recent developments of equipment and techniques have given it a promising new dimension as a retrieval device.

Usually it is possible to couple more than one of these uses of microfilm in any application, but I shall concentrate on the fifth use, as a means of improving retrieval of information.

In considering a microfilm system for storing and retrieving information, it is important to remember that the costs for producing the microfilm and providing viewing equipment are always considerable. Therefore, microfilm is normally justified for information retrieval only when the following conditions exist: (1) The information will be in continuous use for a long period. (2) The information is needed at a single location where there is a high reference rate, or at numerous locations having a moderate to low reference rate.

Roll Microfilm

Roll microfilm is still the predominant type in use today. This is largely due to improved techniques and equipment for producing the film, loading and threading the film in the viewer, making paper enlargements, and locating individual images. Roll microfilm is also receiving increased use for the storage and retrieval of large masses of data produced by computers. Equipment such as the Stromberg-Carlson S-C 4400 Computer Document Recorder, the Burroughs Digitprinter, and the 3M Company Electron Beam Recorder, now make it possible to record information on microfilm direct from computer magnetic tape.

Instead of having to load and thread the microfilm viewer by hand, three manufacturers, the Eastman Kodak Company, the Bell and Howell Corporation, and the 3M Company store the film in cartridges which can be quickly slipped into place in the viewer and the film automatically threaded, ready for instant viewing. All three viewers can be purchased with an attachment for making a paper enlargement of individual film images at speeds of 20 seconds or less.

Four basic methods are used to assist in the rapid location of individual records on the roll of microfilm:

1. Viewers incorporating an odometer-like device, such as those available from the 3M Company and the Bell and Howell Corporation. There are normally about 2 images per inch of microfilm or 2400 linear "locations" per 100 foot roll. At the time of microfilming, a record locator card or sheet is prepared showing the linear location for each

record on the roll of film. When looking up information, the user must determine which roll of microfilm it is on and then refer to the record locator card or sheet to find its linear location on the film. As the film passes through the viewer, he watches the film odometer until the linear location number corresponds with that shown on the record locator card or sheet.

2. Viewers incorporating an image-counting device and keyboard image selection, such as those manufactured by the Eastman Kodak Company. At the time of microfilming, a small opaque square or "blip" is placed beside each film image. A record locator card or sheet is prepared showing the image number for each record on the roll of microfilm. The look-up is much the same as for odometer viewers except that the user merely keys in the image number of "address." A photoelectric cell within the viewer counts the image "blips" and automatically stops at the selected image.

3. Viewers incorporating a numerical or alphabetical scale on the face of the screen, for visual matching with bars or code lines on the microfilm. This technique was introduced by Eastman Kodak, and various adaptations of it are now in widespread use. Small lines or bars representing standard file breaks, such as the 10's, 100's, 1000's positions, are superimposed between page images on the microfilm at the time the records are filmed. The user must have prior knowledge of the identifying name or number of the document he is seeking. Once he has this information, he is able to come to within 10 records or images of the one he is seeking by running the film through the viewer until the bars or code lines match those representing the file break desired on a scale on the viewer screen.

4. Viewers incorporating computer-like circuitry for reading optical binary digital code, such as those available from the National Bureau of Standards; FMA, Inc.; and the Eastman Kodak Company. When microfilming the records, a special camera permits direct recording on the film, in coded form, any desired descriptive data, such as names, numbers, dates, and indexing terms. This descriptive data appears on the film as a pattern of transparent and opaque blocks which form the optical binary digital code. The searching process permits the user, through a keyboard-based device, to locate records on the basis of name, number, etc., and to conduct complex searches involving the coordinate indexing principle.

The nature of admission and registration records, as well as the special information requirements of the office involved, varies enough at each college that any attempt to give specific guidance on the optimum use of microfilm in this paper is fraught with dangers. Each of the tools we have mentioned will be the best answer for some college. But, to give realism to otherwise seemingly sterile facts, let's get down to cases and consider roll microfilm for the permanent record files of the registrar's office. These records appear to fulfill our requirements for using microfilm for information retrieval in that (1) the information will be in continuous use for a long time, and (2) the information is needed at a single location where there is a high reference rate, or at numerous locations having a moderate to low reference rate.

Since our reference need is simply to find records on microfilm by name, there appears little need to go first to a separate index to find the location or proper address on the film such as our first two methods suppose. Since we do not need to use this record to conduct complex searches or use coordinate indexing, nor to search by any feature other than name, the fourth method may be ruled out as unnecessarily expensive. This leaves our third method as the most likely, that of filming our records in alphabetical sequence, with bars or lines on the film that serve as finding aids in much the same way as guide cards in our files.

By this method, retrieval can be rapid, misfiles resulting from removing and replacing records in the file are avoided, and hard copies can be made at the touch of a button. In addition, a copy of the microfilm might also be used as a protection against fire or other disaster. The foregoing exercise in logic is not intended to give the impression that we think roll microfilm is appropriate for the permanent records of all registrars.

The major drawback, as everyone knows, is that the permanent record files do not remain static; they require constant updating. It might be that roll microfilm is not the answer, but before giving it up we should consider how we could reduce this difficulty. In the first place, we would probably not microfilm records of students currently enrolled, thus avoiding the most active records. For the balance, we might splice onto the front of each roll of film microfilm supplements giving the names or updated papers of all students in the alphabetical segment involved whose records have changed during the preceding year. The microfilm user would then scan these additions before proceeding further into the roll. This takes less time than you might suppose. It is the way Social Security does it for employee earning records, except that they splice onto the front of each roll the changes every quarter.

Under this system we might want to re-microfilm the entire file about every 5 to 10 years. This means that we would never destroy the basic hard copy record after microfilming, or at least not until the record is perhaps 50 years old and it is safe to assume it will not be reactivated.

As a parting thought about roll microfilm, the ability to record over 2,000 pages of information on a 100 foot roll of microfilm is not necessarily an advantage. If the look-up or searching pattern is such that the user or system operator needs to retrieve only one or two records per roll, the operation can become excessively time consuming, due to the constant loading and unloading of the rolls of microfilm, and the feeding of the film through the viewer. And, as we have previously noted, roll microfilm can also present problems in situations where individual records must be up-dated, changed, or deleted.

Unitized Microfilm

In order to overcome the problems of roll microfilm, in situations such as that just described, various microfilm unit record

systems have been developed. The most common unitized records are aperture cards, film jackets, microfiche, film chips, and film strips. The unit records used in such systems vary in size and shape and may contain images of as few as 1 or 2 pages of information or as many as 100. Some, such as those employed in superminiaturized systems, contain up to 10,000 pages on 1 sheet of film.

In the past, the records were usually photographed on standard planetary (overhead) roll microfilm cameras, and then cut into short lengths for mounting on unit records. In many of the newer systems, special cameras are employed which permit direct recording of the image on the microfilm unit record.

The following is a brief description of the main types of microfilm unit records.

1. Film jackets represent one attempt to solve problems created by roll microfilm systems in the selective retrieval of, and making additions to, multiple page records. Film jackets consist of two sheets of plastic, affixed to each other in such a way as to provide slots or sleeves for inserting strips of microfilm. A card or opaque strip at the top is provided for writing the identifying information. The entire jacket, with film inside, is placed in the microfilm viewer for reading. Jackets range in size from about 3-by-5 to 5-by-8 inches. Film jackets provide a simple, fairly inexpensive method for unitizing large individual document or case-type records, permit addition of new documents, and protect file from dirt and scratches.

2. Microfiche, or microfilm cards, are plastic sheets with the photo reductions made directly on the sheet. A 4-by-6 inch sheet typically holds 60 images. Recent developments in the field include "superminiaturized cards" in which reduction ratios of up to 300 to 1 are being used and automated retrieval techniques employed.

Microfiche is primarily used as a method for stocking, reproducing, and distributing lengthy records to many users where the volume is not sufficient to justify using conventional printing techniques.

Microfiche unit records are more expensive to produce than most other forms of microfilm and they do not appear to have as much potential for automated retrieval as some of the others.

3. Aperture cards provide for individual document images to be mounted in cut-out windows or apertures in electrical accounting machine punched cards or edge-notched cards. The document number and other selected data are often coded into the card to permit mechanical sorting and sequence-checking later. In many systems, the original or master aperture card is never loaned; instead, the requester is given a duplicate card or enlarged paper copy of the document. Cards with microfilm apertures are best suited to situations where the individual records have few pages. It used to be that only one microfilm image per card was practical, but today some aperture cards will hold as many as eight images. They permit rapid, relatively low cost duplication to meet needs for widespread distribution of documents and heavy user requests for copies.

4. Microfilm chip systems get their name from the fact that they are little film units cut from roll film and are usually quite small. They are designed primarily to meet the need for extremely rapid retrieval of individual short documents (1 to 3 pages) from large, heavily used files. They also satisfy the possible requirement for associating additional pages with the original document and for deleting individual documents from the file. In most such systems, the record number and a small amount of other descriptive data are added to the document in coded form at the time of microfilming. The chips may be manipulated, sorted, and selected by machines employing photoelectric cells, simple electronic circuits, keyboard selector devices, and a built-in viewing screen. Microfilm chip systems involve very high equipment costs.

5. Microfilm strip systems utilize roll microfilm. They are of three types: (a) strips maintained in small containers with a label showing the document number; (b) strips attached to one side of a card or sheet on which is typed or printed identifying information, including possibly a summary or abstract; and, (c) strips mounted on rigid "sticks" about 12 inches long. The first type of strip system has proved useful for long term storage of lengthy documents having a very low reference activity and need for "on-demand" paper enlargements. The second type of strips provides an effective low cost method for distributing documents of moderate lengths where a summary or abstract would be required in any event but the users have only an occasional need to view the entire document. When such need occurs, the full document is readily accessible but yet very little storage space is required for its storage. The third type would normally be best suited to situations where there is a high reference rate and a need for very fast retrieval. One such system, developed by the Eastman Kodak Company, can be used for rapid look-up of student numbers, post office zip codes, and similar directory or catalog-type data.

In the foregoing discussion of unitized microfilm systems, I have given only a sketchy outline of the considerations involved, enough, I hope, to show that there are many possible ways of overcoming the limitations of roll microfilm.

Once again, let's attempt to relate unitized microfilm to the permanent records without appearing to recommend any system for your specific needs. Several unitized systems deserve your consideration, depending on your special requirements; but for most registrars offices we could probably narrow the choice down to film jackets versus aperture cards because they are: (a) Less expensive than microfiche, chip systems, and strips mounted on sticks. (b) Easier to update than microfiche and strips attached to cards or sheets. (c) Easier to use than strips maintained in small containers and strips attached to cards. (d) Adequate for the retrieval speed required. (e) Adequate to insure finding the record desired, since there is little apparent need for retrieval by any characteristic other than name.

Selection between film jackets and aperture cards might be determined by the number of pages in the permanent records for each person. If it were always limited to one page, the aperture card would almost

certainly prevail. If it were frequently over eight pages, the film jacket would likely win. If it is somewhere in between, other considerations would be controlling, including frequency and ease of updating and the need for mechanical sorting to insure the integrity of the file.

COMPUTER STORAGE AND RETRIEVAL

Microfilm information storage and retrieval systems are primarily used as a direct substitute for hard copy storage and retrieval and usually where the look-up is on the basis of name, number, or some other single identifier. Some of the more costly systems, however, will permit searching by several identifiers. Microfilm usually is much less expensive than the computer for the storage of pages of information. The computer stands supreme when there is a need to manipulate data. Hence, it is usually the information retrieval device preferred for extracting and assembling precise answers and facts from a mass of information, for example, in producing statistical reports and various listings. Because of its ability to erase and update, it is unbeatable in most dynamic information situations in which the most current information is needed as changes occur, such as class registration information during the registration period.

Retrieval Speed

A computer is not necessarily faster than hard copy or microfilm retrieval, unless there is also need for data manipulation. Even random access does not insure faster retrieval for the user if the steps he must take to query the machine are difficult.

How much speed do you need? That depends on the use, of course. For most listings and reports, overnight or a day or two delay would be acceptable. But this is not so for student enrollment information during registration, in the thinking of some collegiate officers using or planning to use remote input-output devices. It is not so for any computer retrieval system that hopes to compete with hard copy and microfilm systems where no data manipulation is involved.

For the usual listings and reports, the comparatively slow computers with tape drives will do.

Random access disk and drum storage should be considered when you need within seconds information about a specific student, class, school, etc.

To store in the computer memory all the information that might be needed by your office, so that it would be available instantaneously upon demand, is much too costly for consideration. To illustrate, a 20K computer could store in memory and give you immediate access to only about 10 pages of narrative or numerical data. Therefore, the internal memory must remain as a temporary holding device with the reservoir of data stored on tapes, disks, or drums.

Looking ahead, the most promising computer device on the horizon to give you ready access speeds to tremendous quantities of data without the delays necessary for tape drives is the "optical memory." The data, in optical binary coded form, is recorded directly onto photosensitive disks, chips, or other media by means of electron laser-type beams. In one system, a single 9-inch photo film disk can store the same amount of data which now requires two reels of magnetic tape. Consequently, it makes it possible for the computer to have random access to an almost limitless volume of stored data. Systems having a storage capacity of several trillion bits are already in the final stages of development. Very likely the cost will eventually come within the price range that your institution can afford and you may then have immediate access to all the information you desire to record in machine language.

Two companies, IBM and ITEK, are offering "optical memory" systems. All major computer companies are believed to be developing their own. At present the cost is great, reportedly one-half million dollars for the equipment needed in one system, not including a computer.

All of this adds up to the fact that today you can now have access within seconds to limited volumes of information--if you can afford the price. Tomorrow you can expect access within seconds to almost unlimited volumes of information--at a price you should all be able to afford.

Costs

The most expensive part of any computer system is usually not the equipment, and generally not the cost of information processing and retrieval, but the cost of input. The cost of converting the human language into the machine language understood by the computer is so great that information retrieval in the office environment must generally ride piggy-back on data processing and the accomplishment of the basic tasks of the office.

No doubt when optical scanners have been further perfected, the lowered input costs will change the picture. Full fledged data banks, wherein reposes all student information now kept on paper, may then become a practical reality. For now, any computer retrieval system must be primarily dependent upon information used to process applications, prepare class schedules, issue grade cards, furnish critical administrative reports, and perform other housekeeping chores.

This leads us to the obvious conclusion: Do not buy computers for information retrieval! This does not mean that information retrieval needs should be ignored. Such needs are a part of the total information needs of a university and as such should influence the choice of computers.

Anticipating Needs

You have heard it said that the trouble with computers is that when you install one everybody asks for information they never got

before. This may be a justifiable criticism in some specific instances, but I prefer to believe that officials have needed more and better information all along. Unfortunately, such information vacuums are generally hard to recognize at the outset, possibly because of habit thinking developed in the old environment.

Let me illustrate from the experience of one university I visited. A computer center was established to handle some of the administrative work. Once the fundamental processing jobs had been computerized, they began to get requests for such information as:

1. Names and present addresses of foreign students, by country.

2. Names and addresses of alumni related to present students.

3. Selected information from high school records and aptitude tests to be used in freshman counseling.

4. Student religious affiliation information for use by chaplains.

5. Evaluations of high schools based upon how well students from each school have done while enrolled in this university, for appraising future applicants.

6. Names and addresses of all applicants expecting to major in mathematics, for mailing out information about a new program.

7. Class profile information.

From another university, I received a list of the data outputs of their computer and found 37 that appeared to be derived from admissions and registration data. The list included only recurring information requirements and, I do not know how many special one-time demands are met. Whatever the actual total requirements, it is safe to say that it will increase.

How can you anticipate your needs in order to appraise your systems requirements? The best advice I can give is to examine the experience of universities that have used computers for some time and then allow considerable room for expansion. The thirst for information is virtually insatiable and is limited only by imagination and cost.

File Integrity

Computer filing of information is probably the most accurate filing method we have today. The computer files as accurately as the directions we give it, and the data remain untouched by fallible human hands. This does not mean, however, that we can find the information we want when we want it, even if we have correctly filed it away by computer. That depends upon our ability to retrace our steps by using at least some of the same words or numbers we used when we instructed the computer to file. If we cannot do this, the information will be lost.

Loss of information on tape through aging or accidental erasing is not likely to be a significant consideration if reasonable care is exercised.

Outputs

This brings us to the subject of outputs from the computer. Usually we get the information we want as print-outs. But are you aware that the information can be displayed on a cathode ray tube in human language and almost instantly microfilmed? We have had this for a few years and now someone has improved upon it. A new electron beam process eliminates the cathode-ray tube and records directly on microfilm from the computer. For those of you who are using the computer to prepare and update the permanent record, either of these processes might be of interest. Cost is again our biggest deterrent.

Remote Query

One of the most promising information retrieval developments for your purposes is probably the remote query devices on-line with the computer. The first large scale use of this technique may have been in reserving passenger space on commercial airlines. It provided on-line information, making it possible to accept reservations for involved itineraries right up to flight time.

I was delighted to learn that at least 24 of your offices have indicated current or planned uses of remote input-output equipment in the registration process. A control computer is linked to a number of typewriter or teletype devices from which the machine language information repository may be queried or updated. The computer response is either typed out or appears on a cathode-ray tube in a matter of seconds.

One university reports that in its first experience with this system, the staff was able to register nearly 13,000 students in an average of 3 minutes per student as compared with nearly 50 minutes with the old field house, mass registration method. What is more, the staff believed that the byproducts of computer registration may prove even more significant than the direct benefits of this method of registering.

Time Sharing

Allied to the remote input-output devices is computer time sharing. I am not referring to the possible use of a computer when no one else is using it. I mean the simultaneous use of the computer by several people for different purposes. One remote station might be querying the computer about registration information, another might be asking about classroom space assignments, and a third might be requesting grade information. While all this is going on, the same computer might be preparing a listing or turning out the faculty payroll.

This can be done now but is primarily dependent on a huge memory capacity. Your institution will almost certainly have it within a few years as the cost of large internal storage comes within reach.

In a number of major cities it is possible now to rent access time to such a computer and use it from an input-output station in your office. It may be that for some purposes this will prove less expensive than purchasing or renting your own computer.

Punched Cards

Much of what we have said about computer storage and retrieval is equally applicable to electrical accounting machine punched card systems, except for speed and the advantages made possible by computer memories.

VIDEOTAPE STORAGE AND RETRIEVAL

Some progress has been made in attempting to adapt videotape to office information-retrieval needs. It has the obvious advantage of image storage, plus machine language retrieval, plus the ability to erase and update electronically. Thus, it has many of the advantages of both computer and microfilm systems.

As of now I know of only one actual installation, so our experience is too limited for a satisfactory appraisal.

It is very expensive, as you might suspect, although prices have not been published. The image it will produce on a screen is not as good as that of microfilm. For the present, I doubt that videotape will be used to solve your retrieval problems; but do not rule it out of your future.

STORAGE AND RETRIEVAL WITH COORDINATE INDEXING

No discussion of information storage and retrieval would be complete without considering "coordinate indexing." It is considered by most persons acquainted with the subject as nearly synonymous with "information retrieval." We have skirted around it several times in this paper.

The concept is quite simple when considered in connection with student records. Suppose someone asked you for a roster of those students who had taken certain courses in nuclear physics and mathematics and chemistry and who can speak German but not including any foreign students from Germany. Coordinate indexing makes it possible to furnish such information, though it is virtually impossible to do so by conventional methods.

As another example, suppose you wanted to find a document dealing with the costs for programming computers for retrieval of literature about education. If you were to search a coordinate index under "costs" you might get more documents than you could look through in a week, but by searching under both "costs" and "computers" you might get a more practical number, perhaps 50 documents. Add "computers," "retrieval," "literature," and "education" and you might get precisely the article you want--or nothing at all. Coordinate indexing allows you to broaden or narrow the search to fit your needs.

Computers

Computers can perform coordinate index searching easily but it can be quite involved program-wise. It usually requires an organized list or "dictionary" which give the terms used in the system and defines their meaning. Developing such a dictionary may require considerable skill and time.

In document or literature-type retrieval systems, the computer serves as the index file, compares the terms or "descriptors" you feed into it with the descriptive information of papers or articles it has filed away and gives you the "address" where you can find the items. The items are usually filed by name or number.

In some systems, abstracts or complete text documents are stored on tapes and this is printed out or displayed by the computer rather than furnishing a reference only.

Columnar Cards

Under the simplest coordinate indexing system (the columnar card), one card is used for each indexing term. The first column is used for manually posting serial numbers of relevant documents ending in the digit "0"; the second column is used for those ending in "1," and so forth. In searching, the term cards which apply are visually scanned by first matching the "0" columns of all cards, then the "1" columns, etc.

When we find the same document number on all cards, we know that it is the document we want, since all descriptors are applicable.

Edge-Notched Cards

Edge-notched cards, often called needle-sort cards, have areas around the borders for recording data in a notched code form. The punching positions, represented by pilot holes along one or more of the outer edges of the cards, are reserved for recording such identifying data as keywords or descriptors, dates, and names. The data are recorded on the card by punching out the proper pilot holes in the deck. The selected cards (those which have been notched) fall out, while the others remain on the needle. This process is repeated until

the search is narrowed down to those cards which completely satisfy the search.

Edge-notched card systems are relatively inexpensive and simple to use. Retrieval speeds are satisfactory if the file is small (under 2000 cards) and it is not necessary to incorporate complex indexing and searching techniques. One of the main advantages to the user is immediate availability of pertinent details regarding the document identified through the search. Searching becomes increasingly tedious and time consuming, however, as the file grows beyond 2,000 cards.

Optical Coincidence Cards

The two most common optical coincidence systems, Jonkers Business Machines' "Termatrex" and Royal McBee's "Keydex," employ cards approximately 10-by-11 inches, with positions for punching 10,000 holes. Hand-punched electrical accounting machine punched cards and other types of cards are also sometimes used.

A separate card is prepared for each indexing term used in the system. Each such term card contains a fixed number of assigned but unpunched hole positions. Each of these hole positions represents the same serial number in each card. The documents being indexed are assigned corresponding serial numbers. To enter a document into the system, the term cards which apply are punched in the position reserved for the serial number of that document. When conducting a search, the term cards which apply are superimposed one on top of the other and visually scanned for coinciding holes to identify those documents or things which satisfy the search request. This optical coincidence feature has been dubbed "peek-a-boo."

Optical coincidence cards provide a simple, inexpensive tool for recording and manipulating index data. They are easy to transport and reproduce and can be converted to a computer system. One medical school uses them for various student selection purposes.

CONCLUSION

There is no pat answer to the question: "What information retrieval system is best today?" That depends on your needs, which only you know.

No one type of information retrieval system holds a monopoly on the future, although it does appear that the trend is toward automation and miniaturization. Certainly the computer will figure in your information-retrieval system some place. But somewhere along that information-retrieval spectrum is the right combination for your needs.

ACKNOWLEDGEMENTS

Messrs. Robert Cain and Robert Lando of NARS, GSA, gave considerable assistance in developing and critiqueing this paper. I am also indebted to the following members of AACRAO who provided special information on procedures at their universities.

Mrs. Josephine E. Cortese - American University
Mr. Charles L. Foreman - University of Denver
Mr. Warren R. Haffner - Pennsylvania State University
Mr. Robert T. Ogilvie - American University
Mr. Donald E. Rhoades - University of Iowa
Mr. Don Scherer - Indiana University
Mr. William E. Slaby - Wayne State University